APPLICATION FOR UNITED STATES PATENT

FILTER SET FOR FREQUENCY ANALYSIS

Assignee: Applied Neurosystems Corporation
Attorney Docket No. ANSCP006

By Inventors:

Lloyd Watts
500 Chiquita Avenue #2
Mountain View, CA 94041
A Citizen of Canada

FILTER SET FOR FREQUENCY ANALYSIS

Cross reference to related applications

This application claims priority to co-pending U.S. Patent Application No.

09/534,682 (Attorney Docket No. ANSCP001) entitled EFFICIENT COMPUTATION

5   OF LOG-FREQUENCY-SCALE DIGITAL FILTER CASCADE filed March 24, 2000,

which is incorporated herein by reference for all purposes.

Field of the Invention

The present invention relates generally to signal processing. A system and

method for analyzing a signal into frequency components is disclosed.

10   ## Background Of The Invention

A useful step in analyzing a signal is the separation of the signal into frequency

components. For some time, the fast Fourier transform or FFT algorithm has been used

to analyze a time domain signal into its frequency components. For various types of

processing, and in particular for processing audio signals, it would be desirable to analyze

15   a signal into its frequency components with improved temporal resolution at high

frequencies and better spectral resolution at low frequencies. Numerous techniques have

been proposed for accomplishing this. Included among such techniques are systems that

use a set of filters to separate the signal being analyzed into different channels or

frequency components. Such filter sets operate roughly in a manner that is analogous to

a biological cochlea, which includes a series of filtered output signals that correspond to different frequency channels.

Filter sets may be implemented with analog or digital filters. Previous instantiations of filter sets have been limited by practical considerations in designing

5 filters. For example, high order bandpass filters to separate each channel output are expensive to implement. Various approaches have been implemented using combinations of high pass and low pass filters; however, more efficient techniques are needed to allow real time processing of signals for various important applications including speech recognition, source separation of audio signals and stream separation of

10 audio signals.

Brief Description Of The Drawings

The present invention will be readily understood by the following detailed description in conjunction with the accompanying drawings, wherein like reference numerals designate like structural elements, and in which:

5   Figure 1 is a block diagram illustrating a filter network used in one embodiment for analyzing an input signal into a plurality of frequency components.

Figure 2 is a diagram illustrating an alternative embodiment wherein the low pass filters are not chained together at their inputs and outputs.

Figure 3 is a signal flow graph of a filter equation.

10   Figure 4 is a block diagram illustrating the arrangement of the filters.

Figure 5 is a diagram illustrating an example of the filter response of a second-order section with poles only.

Figure 6 is a diagram illustrating a typical filter response where Q_p is the Q of the pole, Q_z is the Q of the zero, f_{cp} is the center frequency of the pole (also referred to as f_p),

15   and f_{cz} is the center frequency of the zero (also referred to as f_z).

Figure 7 is a diagram illustrating filter responses for filters designed according to the critical band.

Figure 8 is a diagram illustrating the phase characteristics for filters designed according to the critical band.

20   Figure 9A is a diagram illustrating how a filter set as described herein is used in a voice recognition system.

Figure 9B is a diagram illustrating how a filter set as described herein is used in an audio stream separation system.

Figure 9C is a diagram illustrating how a filter set as described herein is used in a spatial correlator or sound localization system.

5

Detailed Description

A detailed description of a preferred embodiment of the invention is provided

below. While the invention is described in conjunction with that preferred embodiment,

it should be understood that the invention is not limited to any one embodiment. On the

5   contrary, the scope of the invention is limited only by the appended claims and the

invention encompasses numerous alternatives, modifications and equivalents. For the

purpose of example, numerous specific details are set forth in the following description in

order to provide a thorough understanding of the present invention. The present

invention may be practiced according to the claims without some or all of these specific

10   details. For the purpose of clarity, technical material that is known in the technical fields

related to the invention has not been described in detail so that the present invention is

not unnecessarily obscured.

A filter cascade for frequency analysis is disclosed that includes a number of

features. In various embodiments, the features are implemented either separately or

15   together. For example, in some embodiments, each frequency component is computed by

subtracting the output of a low pass filter from the input to the filter. In this manner a

bandpass signal is derived. In some embodiments, low pass filters are chained or

cascaded with each filter output being fed to the next filter input in a filter set. The

output of the last filter in the set is downsampled, with the filter set itself collectively

20   acting as a high order antialiasing filter. The downsampled filter set output comprised of

lower frequency components may then be more efficiently processed. Filters in the

cascade may be designed so that the Q of the filters varies with frequency.

U.S. Patent Application No. 09/534,682 which was previously incorporated by

reference (hereinafter, "the 682 application") discloses a digital filter cascade for

frequency analysis. The filters in the cascade are chained together and sets of filters are

separated into octaves with downsampling between octaves. Filter parameters are shared

5   among corresponding filters in different octaves. As described herein, advantages may

be realized if filter parameters are varied among octaves in a manner that varies the Q, or

sharpness of the filters among octaves. In one embodiment, the Q is varied substantially

according to critical bandwidth.

Figure 1 is a block diagram illustrating a filter network used in one embodiment

10   for analyzing an input signal into a plurality of frequency components. An input signal

100 is fed to a low pass filter (LPF) 102. The output of LPF 102 is subtracted from input

signal 100 by a subtractor 104. The output at node 106 thus represents the difference

between the signal before and after LPF 102. It emphasizes a band or channel of

frequencies above the cutoff frequency of LPF 102 and whatever the upper frequency

15   cutoff of the input signal happens to be. The output of LPF 102 is similarly directed to

the input of LPF 112 and the difference between the input and the output of LFP 112 is

computed by a subtractor 114 and output at node 116. The output at node 116 represents

another frequency channel that emphasizes frequencies between the cutoff frequencies of

LPF 102 and LPF 112. In a similar manner, LPF 122 and LPF 132 and subtractor 124

20   and 134 output other frequency channels at nodes 126 and 136. The output of the nodes

may be further processed as is appropriate. For example, in some embodiments, the

outputs are half wave rectified and in some embodiments, the gain of the outputs is

adjusted to compress or expand the dynamic range.

In different embodiments, second order or higher digital or analog filters may be used. The nature of the filters, of course determines the exact nature of each channel output that generally emphasizes a given frequency band and thus has a general bandpass character. Collectively, the channel outputs represent the frequency components of the signal. Because of the subtraction of each LPF input and output, each channel output represents a band or slice of frequencies and the sum of all the outputs represents the entire input signal.

Because the output of each LPF is fed to the input of the next LPF, forming a chain of low pass filters, the output of the last LPF in the chain has characteristics of a much higher order filter than the order of the last filter. This higher order filtering effect may be exploited when the output of the last filter in the chain is downsampled. Essentially, the chain of low pass filters used to separate out frequency channels collectively act as a high order filter that performs the function of an anti aliasing filter when the signal is downsampled.

An example of this is depicted in Figure 1 where downsampler 140 downsamples the output from LPF 132. It should be noted that only four filters are shown in the chain for the purpose of illustration. In most embodiments, more than four filters would be used to process a frequency range before downsampling. The downsampled signal output from downsampler 140 is then processed by another chain of low pass filters that includes LPF 142, LPF 142, LPF 142, LPF 142 and frequency channel outputs are derived by subtractors 144, 154, 164 and 174 at nodes 146, 156, 166 and 176.

In one embodiment, second order individual filters are used and a chain of 60 filters process one octave of the signal before downsampling. Downsampling may be

implemented by simply discarding every other sample or any other appropriate technique. The amount of downsampling is determined by the Nyquist criterion. A suitable amount of oversampling may be done as desired. The combined effect of the chain of filters is that of a very high order anti aliasing filter. Thus, downsampling the signal may be done

5   to speed the processing of lower frequency octaves without requiring an expensive high order anti aliasing filter.

It should be noted that the benefit of chaining the low pass filters is realized in certain embodiments without implementing the subtractors to calculate the frequency bands. The output of each low pass filter may be used directly to represent the energy in

10   each frequency channel. The output of the last filter in each chain is downsampled with the filter chain itself performing the function of an antialiasing filter.

Figure 2 is a diagram illustrating an alternative embodiment wherein the low pass filters are not chained together at their inputs and outputs. Input signal 200 is fed into low pass filters 202, 204, 206, and 208. The difference between the input and the output

15   of each low pass filter is calculated by subtractors 212, 214, 216 and 218. Again, the differences calculated represent an analysis of the frequency bands or channels of the input signal. However, because the output of each filter is not fed to the input of the next filter, the higher order filter effect in the output of the last filter in the chain described above is not realized.

20   The filter cascade may be implemented using either analog or digital filters. In one embodiment, the filters are implemented as digital filters with cutoff frequencies designed to produce the desired channel resolution. Each filter has a set of coefficients

$(a_0, a_1, a_2, b_1, b_2)$ associated with it. The output of each filter is calculated according to the following function:

Equation 1. $\quad y_n = a_0 x_n + a_1 x_{n-1} + a_2 x_{n-2} - b_1 y_{n-1} - b_2 y_{n-2}$

where the filter output y_n is a function of the input data x_n at time n, previous inputs x_{n-1} and x_{n-2}, and previous outputs y_{n-1} and y_{n-2}. Figure 3 is a block diagram illustrating this signal flow. The output of the filter y_n is passed to the input x_n of the next filter in the cascade.

The filter response $H(z)$ is given by the following:

Equation 2. $\quad H(z) = \dfrac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}}$

and $z = e^{i*(\omega/\omega s)}$, $\omega = 2\pi f$, $\quad \omega_s = 2\pi f_s$

where f_s is the sampling frequency.

Substitution of the above into the transfer function of Equation 2 produces a filter response $H(f)$, which is a function of the filter coefficients a_0, a_1, a_2, b_1, b_2 and the sampling rate f_s.

As described in the 682 application, the filter coefficients may be reused between sets of filters with the response of the filters being altered as a result of downsampling between the sets of filters. In the embodiment shown, the filters are evenly distributed over the octaves, resulting in 60 filters per octave. 60 objects are created in a computer. Each object has a set of coefficients as described above, and additionally has ten sets of state variables, corresponding to ten filters running at frequencies that are whole octaves

apart. The 60 objects using their first sets of state variables correspond to the first octave group of filters, while the 60 objects using their second sets of state variables (and sampling at a lower frequency) correspond to the second octave group of filters, and so on. In another embodiment, each object contains a set of coefficients, but only one set of

5   state variables, and is run at a single frequency. In this case, 600 objects are required to represent 600 filters.

The filters in the first octave are tuned to the frequencies in the highest octave, 20 kHz to 10 kHz, and are sampled at 44.1 kHz, which satisfies the Nyquist sampling criterion. The filters in the second octave are tuned to half of the frequencies of the

10   corresponding filters in the first octave, and range from 10 kHz to 5 kHz. These filters in the second octave are sampled at 22.05 kHz, half of the first sampling frequency. Coefficients for each filter are stored in memory and applied in the computations for the filters. The cascade response is the sum of responses of individual filters (which are all weak responses by themselves, but when summed, produce a much stronger response).

15   The coefficients of the filters are determined by the desired response.

As the audio signal is passed through each filter, the signal is sampled and filtered before being passed to the next filter. Figure 4 is a block diagram illustrating the arrangement of the filters. At the end of the first octave, the signal is passed into the first filter in the next octave, which comprises filters sampling at half the sampling rate of the

20   first octave, as stated above. Successive octaves are downsampled in a similar manner, using the same factor of two. In this configuration, each stage acts as an anti-aliasing filter for later stages, removing the high frequencies sufficiently to allow downsampling without aliasing. No extra anti-aliasing filters are required.

Downsampling each successive octave significantly decreases the computational complexity of the system. In addition, the required precision for filter coefficients is lower, and thus, fewer bits are required to represent each coefficient. Digital low-pass filters have the property that the numerical precision required to represent the filter

5   coefficients depends on the ratio between the cutoff frequency and the sampling frequency. For a given sampling frequency, a filter with a low cutoff frequency will require higher-precision coefficients than a filter with a higher cutoff frequency. Without the successive downsampling technique, very high-precision filter coefficients (on the order of 23 bits) are required to represent the lowest-cutoff-frequency filters (30 Hz) at

10   the 44 kHz sampling rate. With the successive downsampling technique, lower-precision coefficients (on the order of 12 bits) can be used to represent the 30-Hz cutoff filters, since the sampling rate is much lower in the lowest octave after many downsampling steps. This reduced precision results in lower hardware complexity (less memory, smaller registers, lower-precision arithmetic operators) and thus lower overall cost in a

15   custom hardware implementation.

In the embodiment described in the 682 application, each filter shares filter parameters with filters that are one, two, or more octaves higher or lower, resulting in reduced storage requirements. For example, the highest frequency filter 40 in the first octave shares filter coefficients with the highest frequency filter 50 in the second octave,

20   the highest frequency filter 60 in the third octave, and so on. The second-highest frequency filter 42 in the first octave shares filter coefficients with the second-highest frequency filters 52 and 62 in the second and third octaves, and with all other corresponding filters (tuned to frequencies that are one, two, or more octaves lower).

Alternatively, it has been determined that the delay at low frequencies can be

improved by changing the filter parameters within each octave as described below. For

many systems, this is preferable to sharing filter parameters between corresponding filters

in different octaves because the benefit from improved delay at low frequencies offsets

5   increased memory storage requirements.

In one embodiment, filter coefficients are tuned to produce a desired Q (quality

factor, or degree of sharpness or frequency selectivity) depending on the frequency band

(determined by the frequency cutoff) being processed by the filter. Reusing filter

coefficients in the cascade results in a cascade with constant Q, and all the filter

10   responses will have the same shape (Q). This "constant-Q" configuration has the

advantages of conceptual simplicity and shared filter coefficients, but has significant

delays at low frequencies. For example, for a constant-Q design with a phase

accumulation of four cycles at all frequencies, the delay at the 20 kHz tap will be 200 μs,

while the delay at the 20 Hz tap will be 200 ms. Faster performance at low frequencies

15   is desirable to improve the response time of the cascade, which may be accomplished by

changing the filter coefficients of the filters in lower octaves.

Figure 5 is a diagram illustrating an example of the filter response of a second-

order section with poles only. The filter may be described in terms of the time constant

20   Tau and quality factor Q, or in terms of filter coefficients b_1 and b_2 mentioned

previously. Tau is the inverse of the center frequency f_c and describes where the peak is,

while Q describes how sharp the peak is. As can be seen from Figure 5, a higher Q

results in a sharper peak, while f_c indicates where the peak occurs. The equations for the

filter are as follows:

$$Vout(z) = \frac{1}{1 - b_1 z^{-1} - b_2 z^{-2}} \, Vin(z)$$

and

$$Vout(s) = \frac{1}{1 + Tau \, s/Q + Tau^2 s^2} \, Vin(s)$$

where the relationship between Tau, Q and b_1, b_2 are given in the "Lyon's

Cochlear Model" Apple Technical Report #13 by Malcolm Slaney ©1988 which is herein

incorporated by reference. The filter coefficients for the filter can be determined from the

center frequency $f_c = 1/Tau$, and the Q of the filter.

The filters may be designed to have zeros as well as poles, and the equation for

such a system is given by

$$Vout(s) = \frac{1 + Tau_z \, s/Q_z + Tau_z^2 s^2}{1 + Tau_p \, s/Q_p + Tau_p^2 s^2} * Vin(s)$$

Figure 6 is a diagram illustrating a typical filter response where Q_p is the Q of the

pole, Q_z is the Q of the zero, f_{cp} is the center frequency of the pole (also referred to as f_p),

and f_{cz} is the center frequency of the zero (also referred to as f_z). The zeros arrest the

dropping gain, and reverse the phase back up to zero. The closer the zero is to the pole,

the sooner these effects occur. If the zero is very close to the pole, the phase trajectory

may not get very far (a small fraction of a cycle) before the zero reverses it. This

property is the key to controlling the total amount of phase accumulation through the

cascade, and hence the delay response of the cascade.

If 600 filters are used, and implemented with a cascade of 600 poles-only sections, each one would contribute a quarter-cycle of phase accumulation at its best frequency, resulting in a large amount of delay. In one embodiment, the filter cascade is configured so that the center frequencies decrease exponentially through the cascade. The Q's decrease gradually through the cascade, to give sharp responses at high frequencies, where delay is not an issue, and to give fast responses at low frequencies, where some loss of sharpness is acceptable in return for faster response. This implementation of nonconstant Q filters is particularly useful for signal processing systems used, for example in submarine passive sonar, speech recognition, music transcription, audio stream separation and sound localization. It should be noted that this approach is not limited to downsampled filter cascades, and may be used with filter cascades with no downsampling.

Design of a filter cascade with constant-Q involves choosing the range of cutoff frequencies and the number of taps per octave, such as a frequency range of 20 Hz to 20 kHz, 600 taps, 10 octaves (60 taps/octave). This determines f_p for each tap. Fixed values are chosen for Q_p, Q_z, and $f_{ratio} = f_z/f_p$, based on the sharpness and delay desired through the cascade. In one embodiment, values used for a constant-Q design may be $Q_p = 7.0$, $Q_z = 7.5$, and $f_{ratio} = 1.03$. In another embodiment, the values may be $Q_p = 23$, $Q_z = 26$, and $f_{ratio} = 1.01$.

For a variable-Q filter cascade using 600 taps in 10 octaves, one embodiment may employ the following values: $Q_p = 7.0$, $Q_z = 7.0$, and $f_{ratio} = 1.03$, with a sampling rate of

44.1 kHz and 2x oversampling in the highest octave. These values are used for the first

360 taps, and then varied linearly over the next 240 taps to $Q_p = 1.6$, $Q_z = 1.6$, and $f_{ratio} = $

1.1 at tap 600 (the lowest frequency tap). This results in a design with broader filter

responses at low frequencies, but much faster time response.

5

In another embodiment, the Q_p, Q_z, and f_{ratio} parameters are selected to match the

filter responses to appropriate psychophysical critical bandwidth and loudness perception

curves. Critical bandwidth is the tuning width of the filter response curves, within which

signal components can interact with each other. Critical bandwidth curves are given in

10   Rossing, 1982, "The Science of Sound" (Addison-Wesley, Reading, MA), the disclosure

of which is hereby incorporated by reference. The critical bandwidth varies from a little

less than 100 Hz at low frequencies to between two and three musical semitones (12% to

19%) at high frequencies. Loudness perception describes how sensitive the filters are to

different frequencies. For example, the threshold of audibility at 20 Hz is about 65 dB

15   higher than at 1 kHz.

One embodiment of a variable-Q filter cascade uses the following parameters:

Tap 0: $Q_p = 7.0$, $Q_z = 7.0$, $f_{ratio} = 1.03$
Tap 300: $Q_p = 11.0$, $Q_z = 11.0$, $f_{ratio} = 1.03$
20   Tap 360: $Q_p = 9.0$, $Q_z = 11.0$, $f_{ratio} = 1.03$
Tap 600: $Q_p = 1.6$, $Q_z = 1.6$, $f_{ratio} = 1.01$

with linear interpolation of parameters between the specified taps. This piecewise

linear variation of the parameters gives a good fit to the psychophysical critical

25   bandwidth and loudness perception curves. Figure 7 is a diagram illustrating filter

responses for filters designed according to the critical band. The filter responses are

sharp at mid-range frequencies, and very broad at low frequencies, corresponding to the critical bandwidth curve. The filters are more sensitive at mid-range frequencies, and about 65 dB less sensitive at low frequencies, so as to match the loudness perception parameters.

5   Figure 8 is a diagram illustrating the phase characteristics for filters designed according to the critical band. The phase characteristics of the filters are such that there are about two cycles of phase accumulation at mid-to-high frequencies, but much less at low frequencies. This results in a faster response at low frequencies, where it is needed.

          A filter cascade for analyzing a signal into frequency components has been

10   described. In various embodiments, the filter cascade utilizes different techniques to improve temporal resolution at high frequencies and spectral resolution at low frequencies. As a result, each of the disclosed filter cascade embodiments are particularly useful as a component of a voice recognition system. In addition, the filter cascade is useful for audio stream separation and sound localization.

15   Figure 9A is a diagram illustrating how a filter set as described herein is used in a voice recognition system. An audio signal is input to a filter set 902 and the output of the filter set is analyzed by a feature extractor 904. The features are classified by a phoneme classifier 906 that matches features with phonemes included in a phoneme database 908. Words are derived based on the phonemes by a word search block 909 that access a word

20   database 910.

          Figure 9B is a diagram illustrating how a filter set as described herein is used in an audio stream separation system such as is described in United States Patent Application No. 60/300,012 (Attorney Docket No.ANSCP003+) by Lloyd Watts (filed

June 21, 2001,) entitled: ROBUST HEARING SYSTEMS FOR INTELLIGENT

MACHINES which is herein incorporated by reference. An audio signal is input to a

filter set 912 and the output of the filter set is analyzed by a set of feature extractors 914

that extract features. The features are grouped by feature grouping processor 916 into

5   separate streams of associated audio signals.

Figure 9C is a diagram illustrating how a filter set as described herein is used in a

spatial correlator or sound localization system such as is described in United States Patent

Application No. 10/004,141 (Attorney Docket No. ANSCP005) by Lloyd Watts (filed

November 14, 2001) entitled: COMPUTATION OF MULTI-SENSOR TIME DELAYS

10   which is herein incorporated by reference. A right channel audio signal is input to a right

channel filter set 922 and a left channel audio signal is input to a left channel filter set

924. The outputs of the filter sets are correlated by a binaural processor 926 to determine

the time delay between the left and right channel input signals. The direction from which

a sound emanates may be determined from the time delay.

15   Although the foregoing invention has been described in some detail for purposes

of clarity of understanding, it will be apparent that certain changes and modifications may

be practiced within the scope of the appended claims. It should be noted that there are

many alternative ways of implementing both the process and apparatus of the present

invention. Accordingly, the present embodiments are to be considered as illustrative and

20   not restrictive, and the invention is not to be limited to the details given herein, but may

be modified within the scope and equivalents of the appended claims.

WHAT IS CLAIMED IS: